



Figure in file

COPY

JDF-5 DNA polymerase nucleotide sequence: 2331 nucleotides (SEQ ID NO: 1)

ATGATCCTTGACGTTGATTACATCACCGAGAATGGAAAGCCCGTCATCAGGGTCTTCAAGAAGGAGAACGG
CGAGTTCAGGATTGAATACGACCGCGAGTTCGAGCCCTACTTCTACGCGCTCCTCAGGGACGACTCTGCCA
TCGAAGAAATCAAAAAGATAACCGCGGAGAGGCACGGCAGGGTCGTTAAGGTTAAGCGCGCGGAGAAGGTG
AAGAAAAAGTTCTCGGCAGGTCTGTGGAGGTCTGGGTCTCTACTTCACGCACCCGCAGGACGTTCCGGC
AATCCGCGACAAAATAAGGAAGCACCCCGCGGTCTACGACATCTACGAGTACGACATACCCTTCGCCAAGC
GCTACCTCATAGACAAGGGCCTAATCCCGATGGAAGGTGAGGAAGAGCTTAACTCATGTCTTCGACATC
GAGACGCTCTACCACGAGGGGAGAAGAGTTTGGAAACCGGGCCGATTCTGATGATAAGCTACGCCGATGAAAG
CGAGGCGCGCGTGATAACCTGGAAGAAGATCGACCTTCCTTACGTTGAGGTTGTCTCCACCGAGAAGGAGA
TGATTAAGCGCTTCTTGAGGGTCGTTAAGGAGAAGGACCCGGACGTGCTGATAACATAACAACGGCGACAAC
TTCGACTTCGCCTACCTGAAAAAGCGCTGTGAGAAGCTTGGCGTGAGCTTTACCCTCGGGAGGGACGGGAG
CGAGCCGAAGATACAGCGCATGGGGGACAGGTTTGGCGTCGAGGTGAAGGGCAGGGTACACTTCGACCTTT
ATCCAGTCATAAGGCGCACCATAAACCTCCCGACCTACACCTTGAGGCTGTATACGAGGCGGTTTTTCGGC
AAGCCCAAGGAGAAGGTCTACGCCGAGGAGATAGCCACCGCTGGGAGACCGGCGAGGGGCTTGAGAGGGT
CGCGCGCTACTCGATGGAGGACGCGAGGGTTACCTACGAGCTTGGCAGGGAGTTCTTCCCGATGGAGGCC
AGCTTTCCAGGCTCATCGGCCAAGGCCTCTGGGACGTTTCCCGCTCCAGCACCGGCAACCTCGTCGAGTGG
TTCCTCCTAAGGAAGGCCTACGAGAGGAACGAACCTCGCTCCCAACAAGCCCGACGAGAGGGAGCTGGCGAG
GAGAAGGGGGGGCTACgCCGGTGGCTACGTCAAGGAGCCGGAGCGGGGACTGTGGGACAATATCGTGTATC
TAGACTTTTCGTAGTCTCTACCTTCAATCATAATCACCCACAACGTCTCGCCAGATACGCTCAACCGCGAG
GGGTGTAGGAGCTACGACGTTGCCCCGAGGTGCGTCAAGGTTCTGCAAGGACTTCCCGGCTTCATTCC
GAGCCTGCTCGGAAACCTGCTGGAGGAAAGGCAGAAGATAAAGAGGAAGATGAAGGCAACTCTCGACCCGC
TGGAGAAGAATCTCCTCGATTACAGGCAACGCGCCATCAAGATTCTCGCCAACAGCTACTACGGCTACTAC
GGCTATGCCAGGGCAAGATGGTACTGCAGGGAGTGCGCCGAGAGCGTTACGGCATGGGGAAGGGAGTACAT
CGAAATGGTCATCAGAGAGCTTGAGGAAAAGTTCGGTTTTAAAGTCCTCTATGCAGACACAGACGGTCTCC
ATGCCACCATTCTTGAGCGGACGCTGAAACAGTCAAGAAAAAGGCAATGGAGTTCTTAACTATATCAAT
CCCAAACTGCCCCGCCCTTCTCGAACTCGAATACGAGGGCTTCTACGTCAGGGGCTTCTTCGTCACGAAGAA
AAAGTACGCGGTCTACGACGAGGAGGCAAGATAACCACGCGCGGGCTTGAGATAGTCAGGCGCGACTGGA
GCGAGATAGCGAAGGAGACGCGAGGCGAGGGTTTTGGAGGCGATACTCAGGCACGGTGACGTTGAAGAGGCC
GTCAGAATTGTCAGGGAAGTCACCGAAAAGCTGAGCAAGTACGAGGTTCCGCCGAGAGAAGCTGGTTATCCA
CGAGCAGATAACGCGCGAGCTCAAGGACTACAAGGCCACCGGCCCGCACGTAGCCATAGCGAAGcGTTTGG
CCGCCAGAGGTGTTAAATCCGGCCCGGAACGTGTGATAAGCTACATCGTTCTGAAGGGCTCCGGAAGGATA
GGCGACAGGGCGATTCCCTTCGACGAGTTCGACCCGACGAAGCACAAGTACGATGCGGACTACTACATCGA
GAACCAGGTTCTGCCGGCAGTTGAGAGAATCCTCAGGGCCTTCGGCTACCGCAAGGAAGACCTGCGCTACC
AGAAGACGAGGCAGGTGGGCTTGCGCGTGGCTGAAGCCGAAGGGGAAGAAGAAGTGA

FIG. 2.
JDF-3 DNA polymerase amino acid sequence (SEQ ID NO: 2)
Theoretical molecular weight: 90.3 kD

MILDVDYITENGKPVIRVFKKENGFEFRIEYDREFEPYFYALLRDDSAIEEIKKITAERHGRVVKVKRAEKV
KKKFLGRSVEVWVLYFTHPQDVPAIRDKIRKHPAVIDIYEYDIPFAKRYLIDKGLIPMEGEEELKLMSFDI
ETLYHEGEEFGTGPILMISYADESEARVITWKKIDLPYVEVVSTEKEMIKRFLRVVKEKDPDVLITYNGDN
FDFAYLKKRCEKLGVSFTLGRDGSEPKIQRMGDRFAVEVKGRVHFDLYPVIRRTINLPTYTLEAVYEAVFG
KPKEKVYAEIATAWETGEGLEVARYSMEDARVTYELGREFFPMEAQLSRLIGQGLWDVSRSSSTGNLVEW
FLLRKAYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSPDTLNRE
GCRSYDVAPEVGHKFKCDFPGFIPSLLGNNLEERQKIKRKMKATLDPLEKNLLDYRQRAIKILANSYYGYY
GYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYADTDGLHATIPGADAETVKKKAMEFLNYIN
PKLPGLLELEYEGFYVRGFFVTKKKYAVIDEEGKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEA
VRIVREVTEKLSKYEVPPEKLVIEHQITRELKDYKATGPHVAIAKRLAARGVKIRPGTVISYIVLKGSGRI
GDRAIPFDEFDPTKHKYDADYYIENQVLPAYERILRAFGYRKEDLRYQKTRQVGLGAWLKPKGKKK

FIG. 3.

JDF-3 DNA polymerase with intein sequence (SEQ ID NO: 3)

MILDVDYITENGKPVIRVFKKENGFEFRIEYDREFEPYFYALLRDDSIAIEE
IKKITAERHGRVVKVKRAEKVKKKFLGRSVEVWVLYFTHPQDVPAIRDKI
RKHPAVIDIYDYDIPFAKRYLIDKGLIPMEGEEELKLMSFDIETLYHEGE
EFGTGPILMISYADESEARVITWKKIDLPYVEVVSTEKEMIKRFLRVVKE
KDPDVLITYNGDNFDFAYLKKRCEKLGVSFTLGRDGSEPQIQRMGDRFAV
EVKGRVHFDLYPVIRRTINLPTYTLEAVYEAVFGKPKKVKYAEIATAWE
TGEGLERVARYSMEDARVTYELGREFFPMEAQLSRLIGQGLWDVSRSTG
NLVEWFLLRKAYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNI
VYLDFRSLYPSIIITHNVSPDTLNREGCRSYDVAPEVGHKFKCDFPGFIP
SLLGNLLEERQKIKRKMKATLDPLEKNLLDYRQRAIKILAN

Extein 1

SLLPGEWVA
VIEGGKLRPVRIGELVDGLMEASGERVKRDGDTEVLEVEGLYASPSTGSP
RKPAQCR*KP**GTAMPGKfte*LSTPEGGLSVTRGHSLFAYRDASLWR*
RGRRRFKPGDLLAVPSG*PSRRGGRGSTSLNCSSNCPRRKRPTCHRHSGK
GRKNFFRGMRLRTLWIFGEEKTGGRPGATWSTLRGLGYVKLRKIGYGVVD
REGLGKVPFRFYERLVEVIRYNGNRGEFIADFNALRPVLRIMMPEKELEEW
LVGTRNGFRIRPFIEVDWKFAKLLGYVSEGSAGKWKNTGGWSYSVRLY
NEDGSVLDDMERLARSSLGA*ARGELRRDFKEDGLHNLRGALRFTGREQE
GSVAYLHVP*GGPLGLP*GVLHRRRRRSPEQDGSALHQERASG*RPRPAP
ELAGRLSDKRPPRQRLQGLRERGTALYRVPEAEERLTYSHVIPREVLEE
TSAGPSRRT*VTGNSGSWWKAGSSTRKGPVG*AGSSTGI*SSTGSRKSGR
KATRGTTT*ALRRTRTSGGLWVPLRAQX

Intein 1

SYYGYYGYARARWYCRECAES
VTAWGREYIEMVIRELEEKFGFKVLYADTDGLHATIPGADAETVKKKAME
FLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEEGKITTRGLEIVR
RDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKLVI
HEQITRELKDYKATGPHVAIAKRLAARGVKIRPGTVISYIVLKGSGRIGD
RAIPFDEFDPTKHKYDADYYIENQVLPAVERILRAFgyrkEDLRYQKTRQ
VGLGAWLKPKGKKK

Extein 2

FIG. 4.
JDF-3 DNA polymerase genomic sequence (SEQ ID NO: 4)

AATTCCACTGCCGTGTTTAACCTTTCCACCGTTGAACTTGAGGGTGATT
TCTGAGCCTCCTCAATCACTTAATCGAGACCGCGATTACCTTGAACGG
TACACGTTCAACGATTCGGTTCTTGTAATGGTCGATACTGGGCCGTGCTG
GATTTTCTAAACGTCTCAAGAACGGCTTTCATCAACGGAACTGCCACGT 5' untranslated sequence
CTCCGCCGTCTGAGGGTTAAACCTGAAGTTCAAGACTTTGCAACGGAAT
GGCGAGAGAACGGCGACTACCCAGTGAAGAGCTTTTGAAAGCCAAAGC
CGAGCTTCAGCGAATGTGCGGTGCCCTTGTTCAAGAGTTGTGAGCCCTTG
ATTGTTGTTTTCTCCTCTTTCTGATAACATCGATGGCGAAGTTTATTAG
TTCTCAGTTCGATAATCAGGCAGGTGTTGGTC

ATGATCCTTGACGTTGAT
TACATCACCGAGAATGGAAAGCCCGTCATCAGGGTCTTCAAGAAGGAGAA
CGGCGAGTTCAGGATTGAATACGACCGCGAGTTCGAGCCCTACTTCTACG
CGCTCCTCAGGGACGACTCTGCCATCGAAGAAATCAAAAAGATAACCGCG
GAGAGGCACGGCAGGGTCGTTAAGGTTAAGCGCGCGGAGAAGGTGAAGAA
AAAGTTCCTCGGCAGGTCTGTGGAGGTCTGGGTCTCTACTTCACGCACC
CGCAGGACGTTCCGGCAATCCGCGACAAAATAAGGAAGCACCCCGCGGTC
ATCGACATCTACGAGTACGACATACCCTTCGCCAAGCGCTACCTCATAGA
CAAGGGCCTAATCCCGATGGAAGGTGAGGAAGAGCTTAACTCATGTCCT
TCGACATCGAGACGCTCTACCACGAGGGAGAAGAGTTTGGAAACGGGCCG
ATTCTGATGATAAGCTACGCCGATGAAAGCGAGGCGCGCTGATAACCTG
GAAGAAGATCGACCTTCCTTACGTTGAGGTTGTCTCCACCGAGAAGGAGA
TGATTAAGCGCTTCTTGAGGGTCGTTAAGGAGAAGGACCCGGACGTGCTG
ATAACATACAACGGCGACAACCTTCGACTTCGCCTACCTGAAAAAGCGCTG
TGAGAAGCTTGGCGTGAGCTTTACCCTCGGGAGGGACGGGAGCGAGCCGA Extein 1
AGATACAGCGCATGGGGGACAGGTTTGCGGTGAGGTGAAGGGCAGGGTA
CACTTCGACCTTTATCCAGTCATAAGGCGCACCATAAACCTCCCGACCTA
CACCCTTGAGGCTGTATACGAGGCGGTTTTTCGGCAAGCCCAAGGAGAAGG
TCTACGCCGAGGAGATAGCCACCGCCTGGGAGACCGGCGAGGGGCTTGAG
AGGGTCGCGCGCTACTCGATGGAGGACGCGAGGGTTACCTACGAGCTTGG
CAGGGAGTTCTTCCCGATGGAGGCCAGCTTTCAGGCTCATCGGCCAAG
GCCTCTGGGACGTTTCCCGCTCCAGCACCGGCAACCTCGTCGAGTGTTTC
CTCCTAAGGAAGGCCTACGAGAGGAACGAACCTCGCTCCCAACAAGCCCGA
CGAGAGGGAGCTGGCGAGGAGAAGGGGGGGCTACGCCGTTGGCTACGTCA
AGGAGCCGGAGCGGGGACTGTGGGACAATATCGTGTATCTAGACTTTTCGT
AGTCTCTACCCTTCAATCATAATCACCCACAACGTCTCGCCAGATACGCT
CAACCGCGAGGGGTGTAGGAGCTACGACGTTGCCCCCGAGGTCGGTCACA
AGTTCTGCAAGGACTTCCCGGCTTCATTCCGAGCCTGCTCGGAAACCTG
CTGGAGGAAAGGCAGAAGATAAAGAGGAAGATGAAGGCAACTCTCGACCC
GCTGGAGAAGAATCTCCTCGATTACAGGCAACGCGCCATCAAGATTCTCG
CCAAC

AGCCTTCTTCCCGGGAGTGGGTTGCGGTCATTGAAGGGGGGAAA
CTCAGGCCCCGTCCGCATCGGCGAGCTGGTTGATGGACTGATGGAAGCCAG
CGGGGAGAGGGTGAAAAGAGACGGCGACACCGAGGTCCTTGAAGTCGAGG
GGCTTTACGCCTCTCCTTCGACAGGGAGTCCAAGAAAGCCCGCACAATGC
CGGTGAAAGCCGTGATAAGGCACCGCTATGCCGGGAAGTTTACAGAATA
GCTCTCAACTCCGGAAGGAGGATTAAGCGTGACGCGCGGCCACAGCCTCT
TCGCGTACCGGGACGCGAGCTTGTGGAGGTGACGGGGGAGGAGGAGGTTT
AAGCCCGGCGACCTCCTGGCGGTGCCAAGCGGATAACCTTCCCGGAGAGG
Intein 1

AGGGAGAGGCTCAACATCGTTGAACTGCTCCTCGAACTGCCCCGAGGAGGA
 AACGGCCGACATGTTCATCGACATTCCGGCAAGGGTAGAAAGAATTCTTC
 AGGGGAATGCTCAGAACCCTCCGCTGGATTTTCGGGGAGGAGAAGACCGG Intein 1
 AGGGCGGCCAGGCGCTACCTGGAGCACCTTGCGTGGGCTCGGCTACGTGA
 AGCTGAGGAAAATCGGCTACGGGGTGGTTGATAGGGAGGGAAGTGGGAAAG
 GTACCGCGCTTCTACGAGAGGCTCGTGGAGGTAATCCGCTACAACGGCAA
 CAGGGGGGAGTTCATCGCCGATTTCAACGCGCTCCGCCCCGTCCTCCGCC
 TGATGATGCCCCGAGAAGGAGCTTGAAGAGTGGCTCGTTGGGACGAGGAAC
 GGGTTTCAGGATAAGGCCGTTTCATAGAGGTTGATTGGAAGTTCGCAAAGCT
 CCTCGGCTACTACGTGAGCGAGGGGAGCGCCGGGAAGTGGAAAAACCGGA
 CCGGGGGCTGGAGCTACTCGGTGAGGCTTTACAACGAGGACGGGAGCGTT
 CTCGACGACATGGAGAGACTCGCGAGGAGTTCTTTGGGGGCGTGAGCGCG
 GGGGGAACCTACGTGAGATTTCAAAGAAGATGGCCTACATAATCTTCGAG
 GGGCTCTGCGGTTACCCGGCCGAGAACAAAGAGGGTTCCGTGGCTTATCTT
 CACGTCCCCTGAGGAGGTCCGCTGGGCCTTCCTTGAGGGGTACTTCATCG
 GCGACGGCGACGTTACCCGAGCAAGATGGTTCCGGCTCTCCACCAAGAGC
 GAGCTTCTGGCTAACGGCCTCGTCCTGCTCCTGAACTCGCTGGGCGTCTC
 AGCGATAAACGTCCGCCACGACAGCGGGGTTTACAGGGTCTACGTGAACG
 AGGAAGTGCCTTTACAGAGTACCGGAAGCGGAAGAAGCCTCACTTACT
 CCCACGTACATCCGAGGGAAGTGCTGGAGGAGACTTCGGCCGGGCCTTCC
 AGAAGAACATGAGTCACGGGAAATTACGGGAGCTGGTGGAAAGCGGGGAG
 CTCGACGCGGAAAGGGCCGGTAGGATAGGCTGGCTCCTCGACGGGGATAT
 AGTCCTCGACAGGGTCTCGGAAGTCAGGAAGGAAAGCTACGAGGGGTACG
 TCTACGACCTGAGCGTTGAGGAGGACGAGAACTTCTGGCGGGCTTTGGGT
 TCCTCTACGCGCACAAACNN

AGCTACTACGGCTACTACGGCTATGCCAGGG
 CAAGATGGTACTGCAGGGAGTGCGCCGAGAGCGTTACGGCATGGGGAAGG
 GAGTACATCGAAATGGTCATCAGAGAGCTTGAGGAAAAGTTCCGGTTTAA
 AGTCCTCTATGCAGACACAGACGGTCTCCATGCCACCATTCTTGAGCGG
 ACGCTGAAACAGTCAAGAAAAAGGCAATGGAGTTCTTAACTATATCAAT
 CCCAAACTGCCCCGGCCTTCTCGAACTCGAATACGAGGGCTTCTACGTGAG
 GGGCTTCTTTCGTACGAAGAAAAAGTACGCGGTTCATCGACGAGGAGGGCA
 AGATAACCACGCGCGGGCTTGAGATAGTCAGGCGCGACTGGAGCGAGATA
 GCGAAGGAGACGCGAGGCGAGGGTTTTGGAGGCGATACTCAGGCACGGTGA Extein 2
 CGTTGAAGAGGCCGTGAGAATTGTGAGGGAAGTCACCGAAAAGCTGAGCA
 AGTACGAGGTTCCGCCGAGAAAGCTGGTTATCCACGAGCAGATAACGCGC
 GAGCTCAAGGACTACAAGGCCACCGGCCGACGTAGCCATAGCGAAGCG
 TTTGGCCGCCAGAGGTGTTAAATCCGGCCCGAACTGTGATAAGCTACA
 TCGTTCTGAAGGGCTCCGGAAGGATAGGCGACAGGGCGATTCCCTTCGAC
 GAGTTCGACCCGACGAAGCACAAAGTACGATGCGGACTACTACATCGAGAA
 CCAGGTTCTGCCGCGAGTTGAGAGAATCCTCAGGGCCTTCGGCTACCGCA
 AGGAAGACCTGCGCTACCAGAAGACGAGGCAGGTCCGGCTTGCGCGGTGG
 CTGAAGCCGAAGGGGAAGAAGAAGTGA

GGAATTATCTGGTTTCTTTTCCC
 AGCATTAAATGCTTCCGACATTGCCTTATTTATGAACTCCTGTTGTGCC
 TGAGTTTGTGCCAGAAAACAGCCTGTTCTGACGGCGCTTTTCTTGCCAG
 GTCTCTTGAGTTTCGCAAGGGTCTTCTCGACCAGCTCAATGGTCTTGTCG
 TCATTGTTTNNNNNNNNNNNNNNNNNNNNCCCGGGGACTTCATACTGGC
 GGTAATAGACAGGGATTCTTCCTCAAGGACTTCCCGGGAGGCATTGGAG
 TTTTTTGGTGGGGCTTTCACAGGATTGCTCATCTTGTTGGATTCTCGTT
 CGATTGAATCTGTCCACTTGAGGGTGTAGGTCGAGACGGTGGAGCGCGTA

TTCCGGGAGCGGGTCTTGAGGCTCCATTTTTTCAGTCCTCCTCCGGCGAAG 3' Untranslated sequence
AAGTGGAACCTCAAGCCGGGTGTTAGCTTATGTTATGTTCCCAACTCCTCC
AGCACCTCCAGGATCCCCCTCAATCCCGGAACCTCGAAGCCCCCTCTCGTGG
ATCTTTCTAACTTCCTCTGCCTCCGGGTTTATCCAGACCGCCACATGCC
GGCTCTCAGCGCACCTCGAAATCCTCCGCGTAGGTGTCGCCGATGTGGA
TTGCCTCGTCCGGCTCGACCCCGAAGCATCGAGCGGTTTTCTGAACATCT
CGGGCATCGGCTTATACGCCAGAACCCTCGTCGGCGAAGAAGGTTCCCTCA
ATGTAGTCCATCAGGCCGAACCTCTCGAGGGGGGGCCCGGTACCCAATTC
GCCCTATAGTGAGTCGATTACAATTCAGTGGCCGTCGTTTTACAACGTCTG
TGACTGGGAAAACCCTGGCGTTACCCAACCTTAAGTCGCTTTGCAGCACAT
CCCCC

BEST AVAILABLE COPY

Preliminary Qualification of Mutants

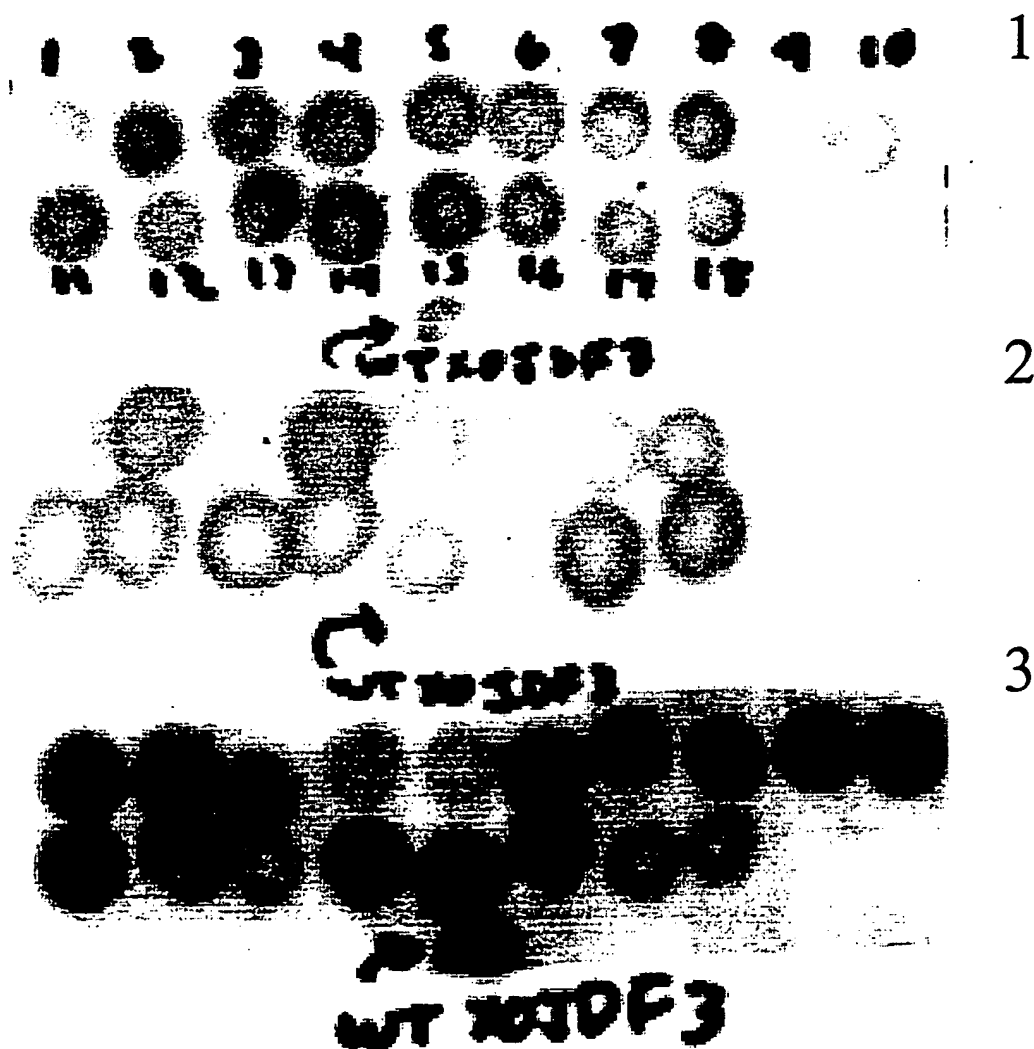


Figure 5

BEST AVAILABLE COPY

Sequencing with Purified Mutants

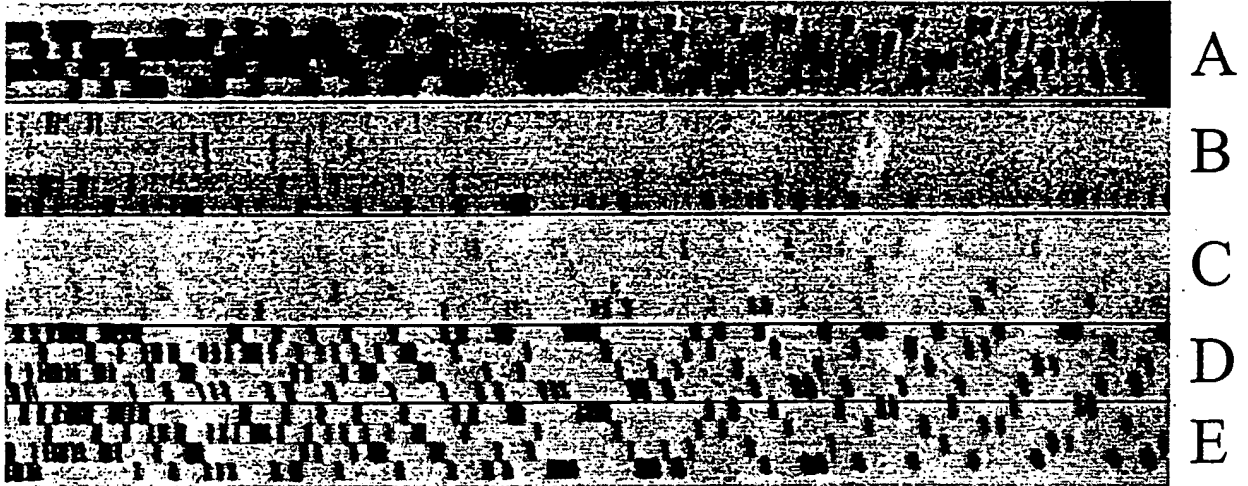


Figure 6

BEST AVAILABLE COPY

Sequencing with Dye-labeled Dideoxynucleotides

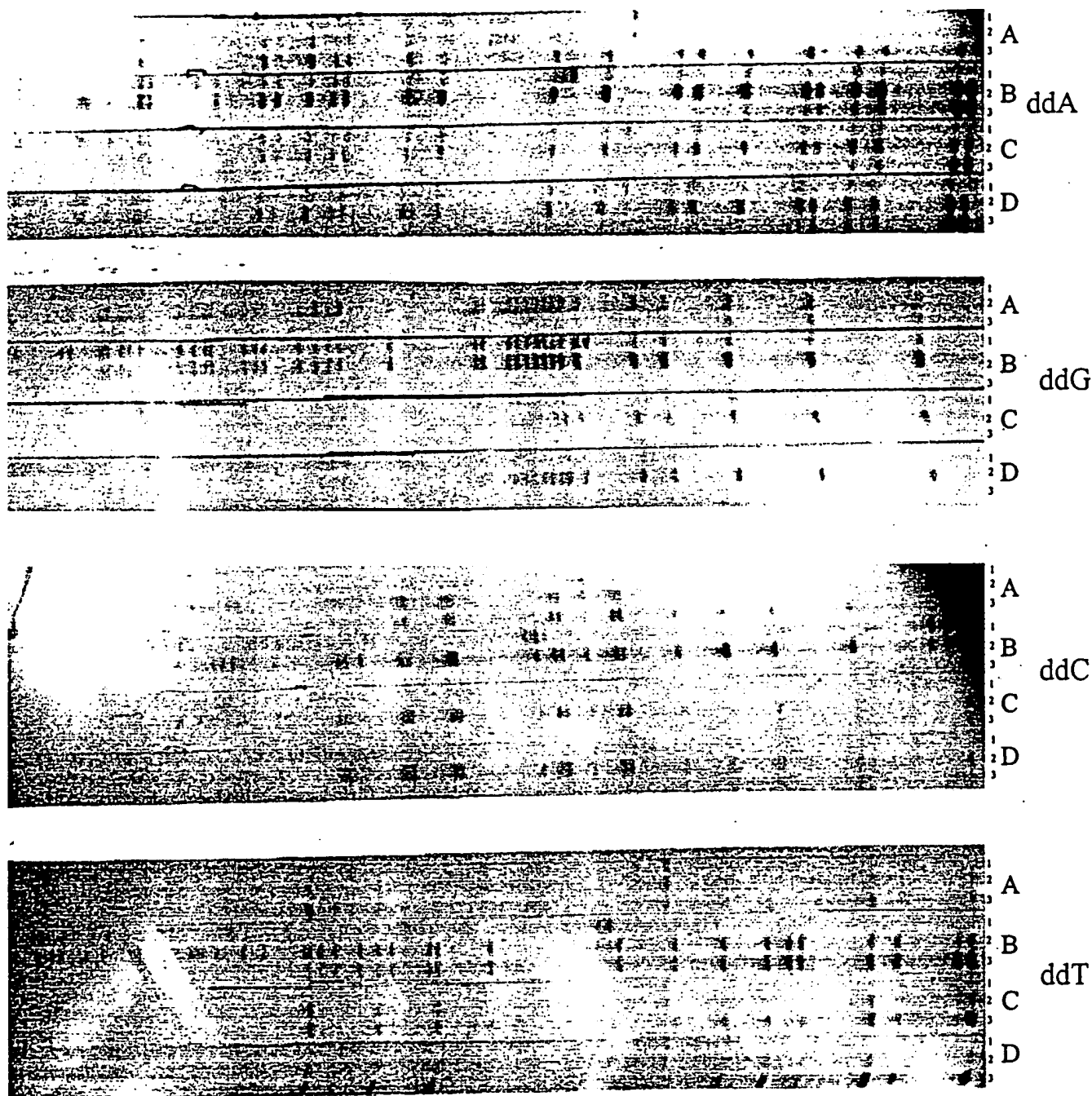


Figure 7

Sequencing with the P410L, A485T Double Mutant and α -³³P Dideoxynucleotides

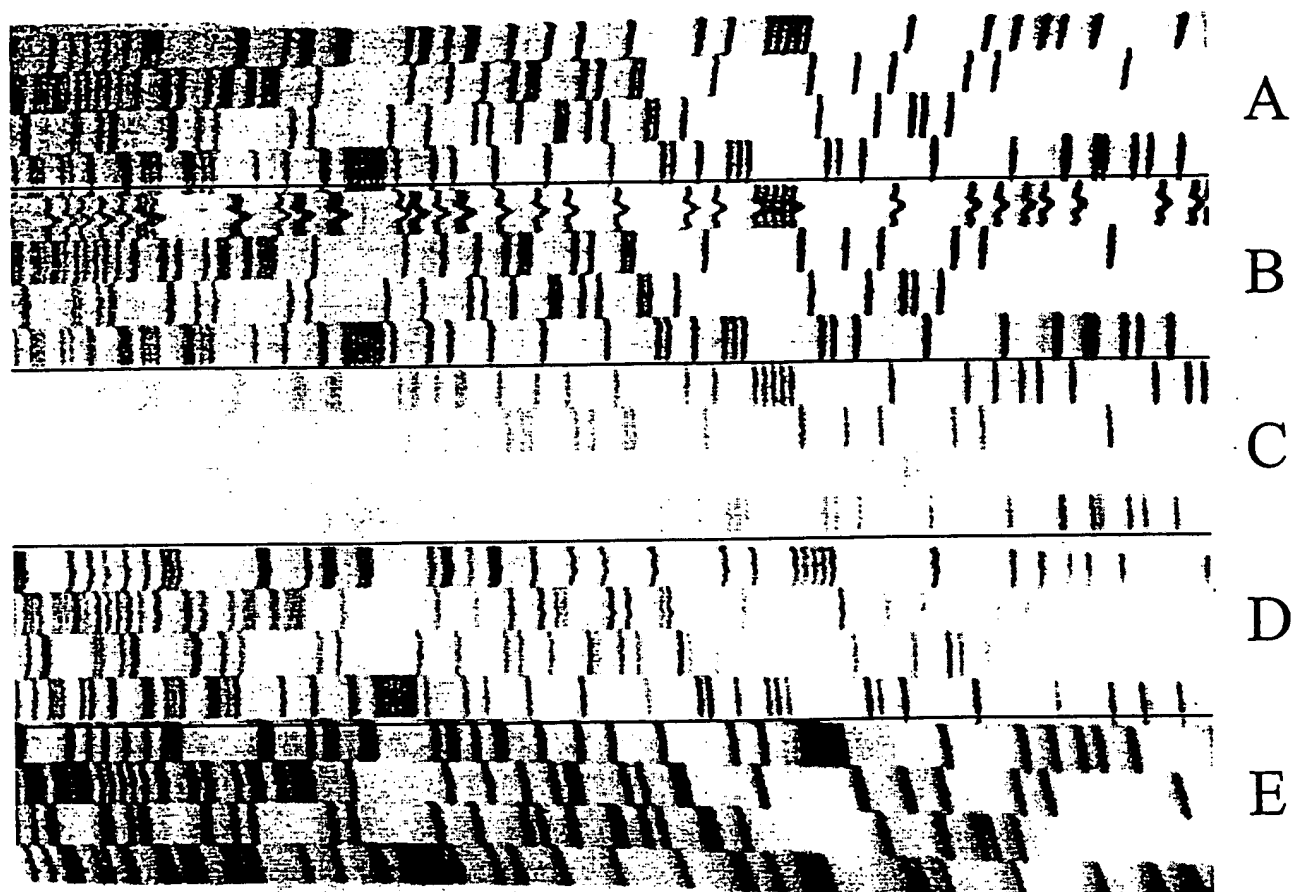
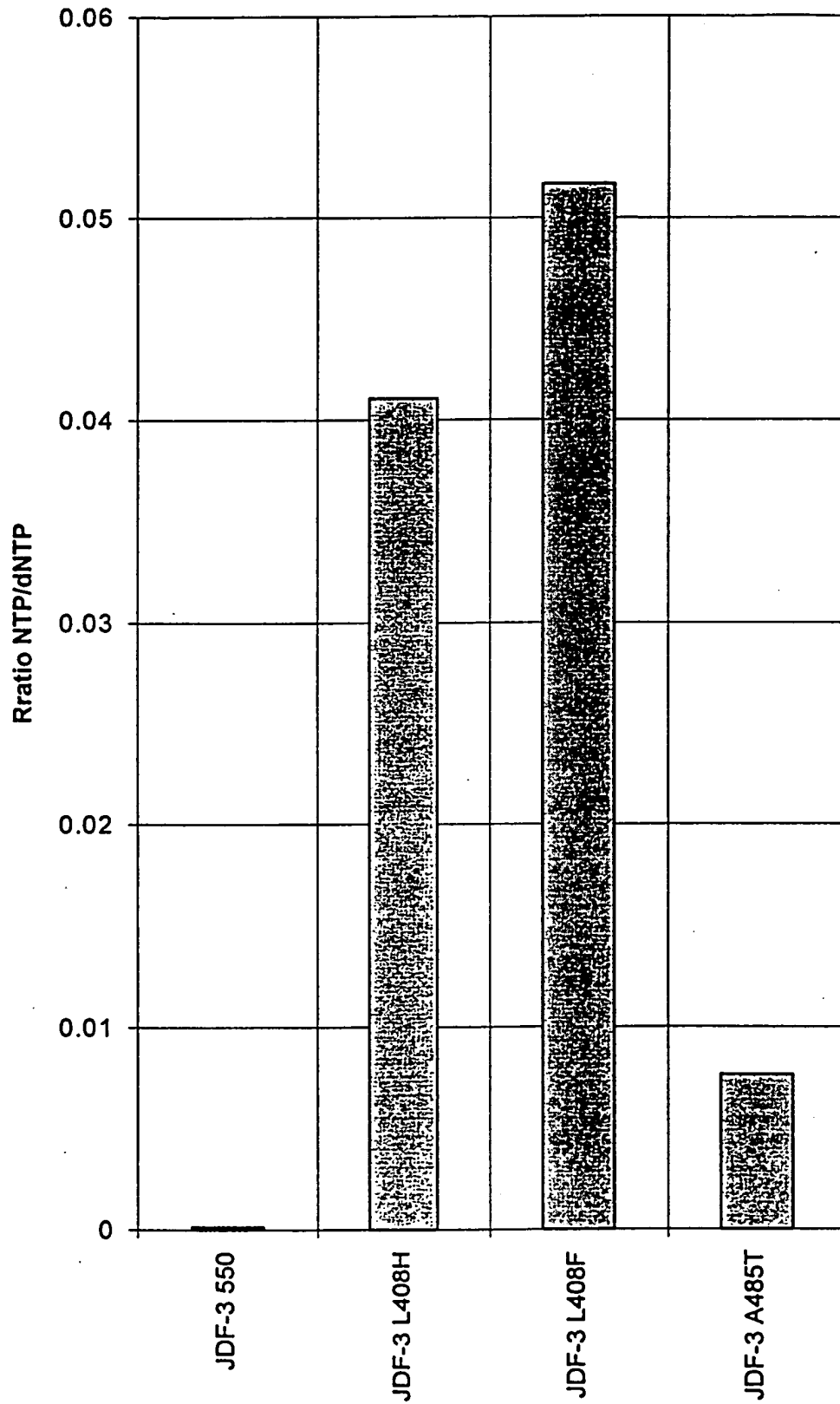


Figure 8

Figure 9

2-29-2007 10:03 AM COPY



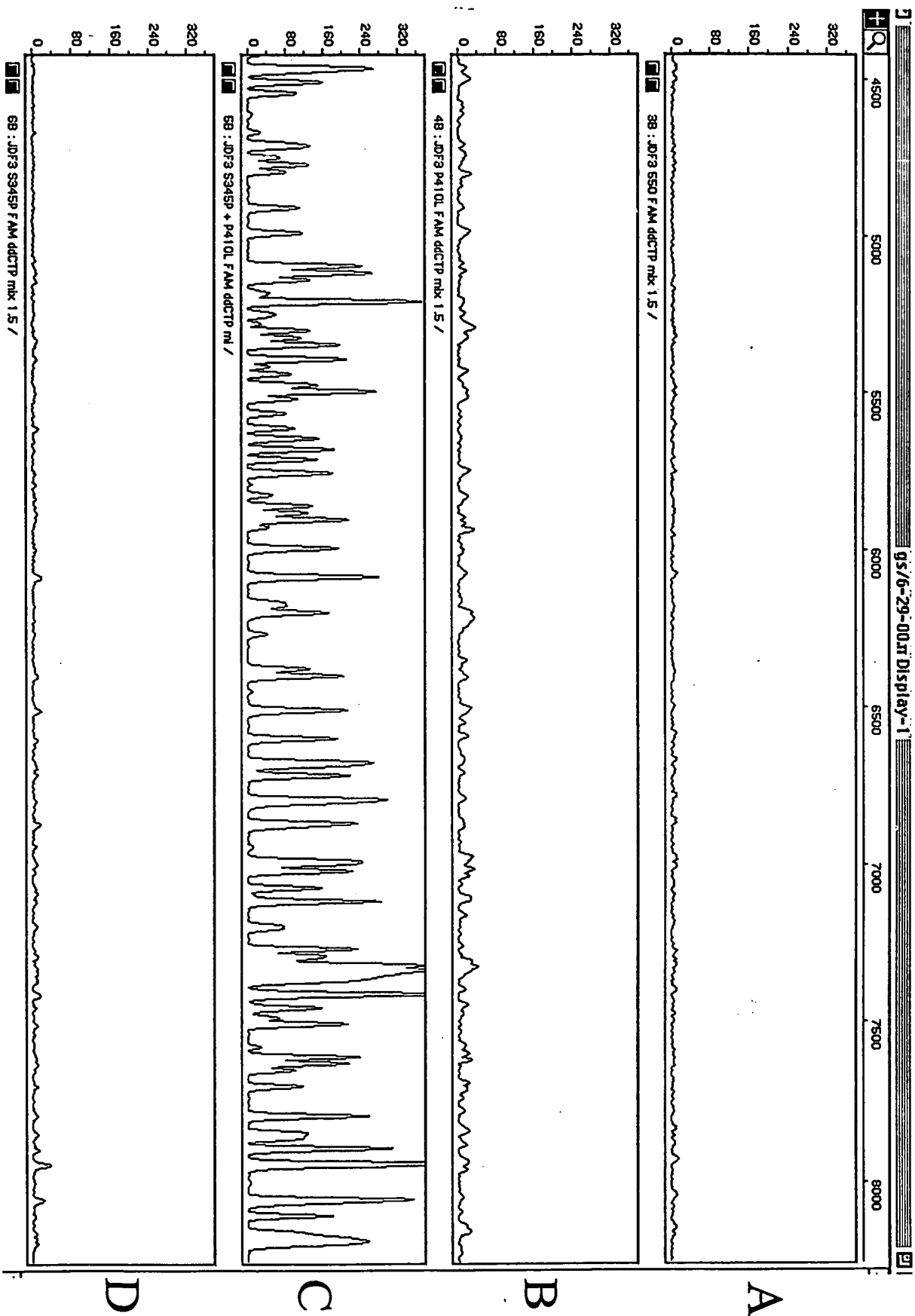
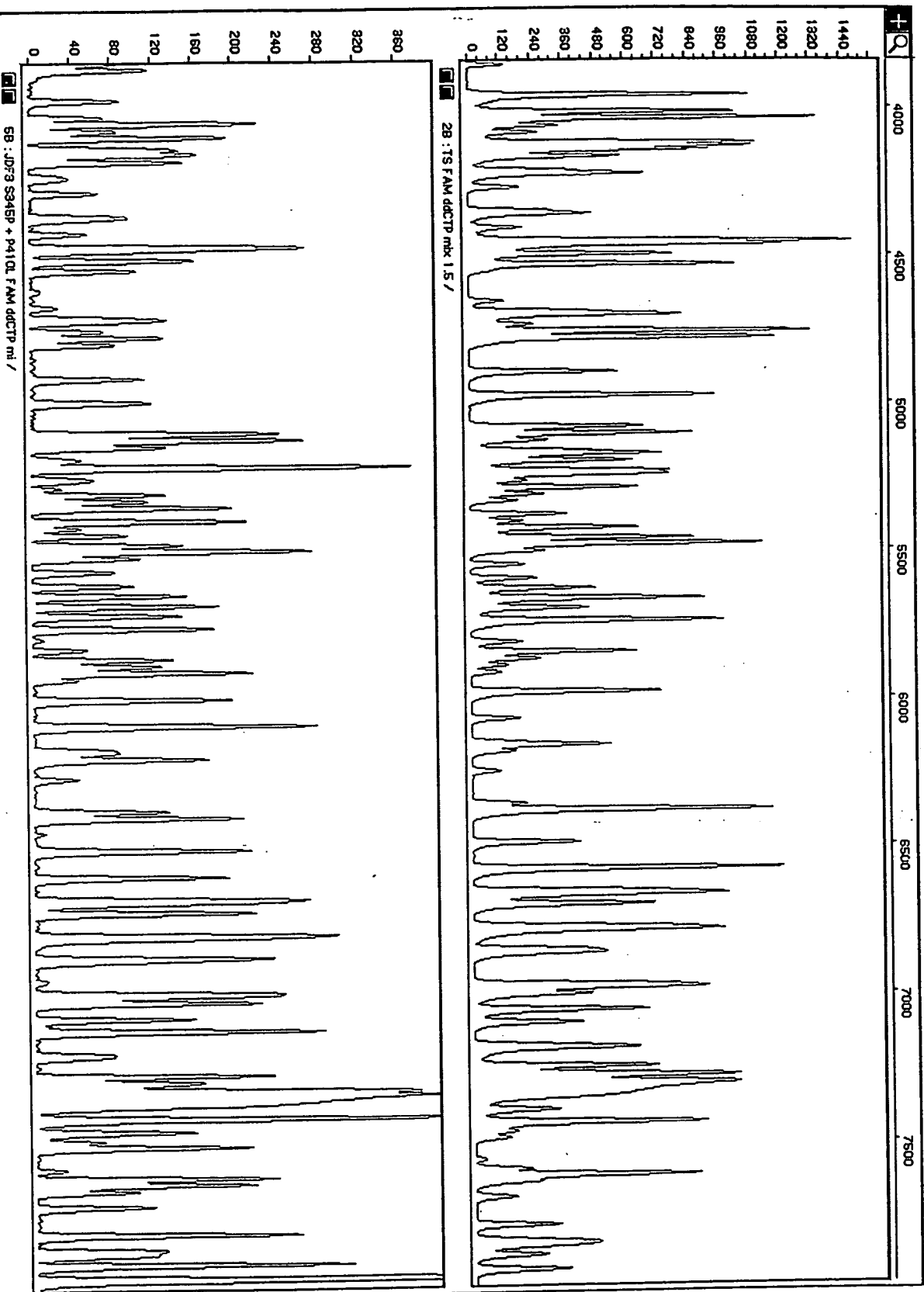


Figure 10



33P-labeled DNA template

^{33}P - TAACGTTGGGGGGGGGGCA →
TGCAACCCCCCCCCCGTAT

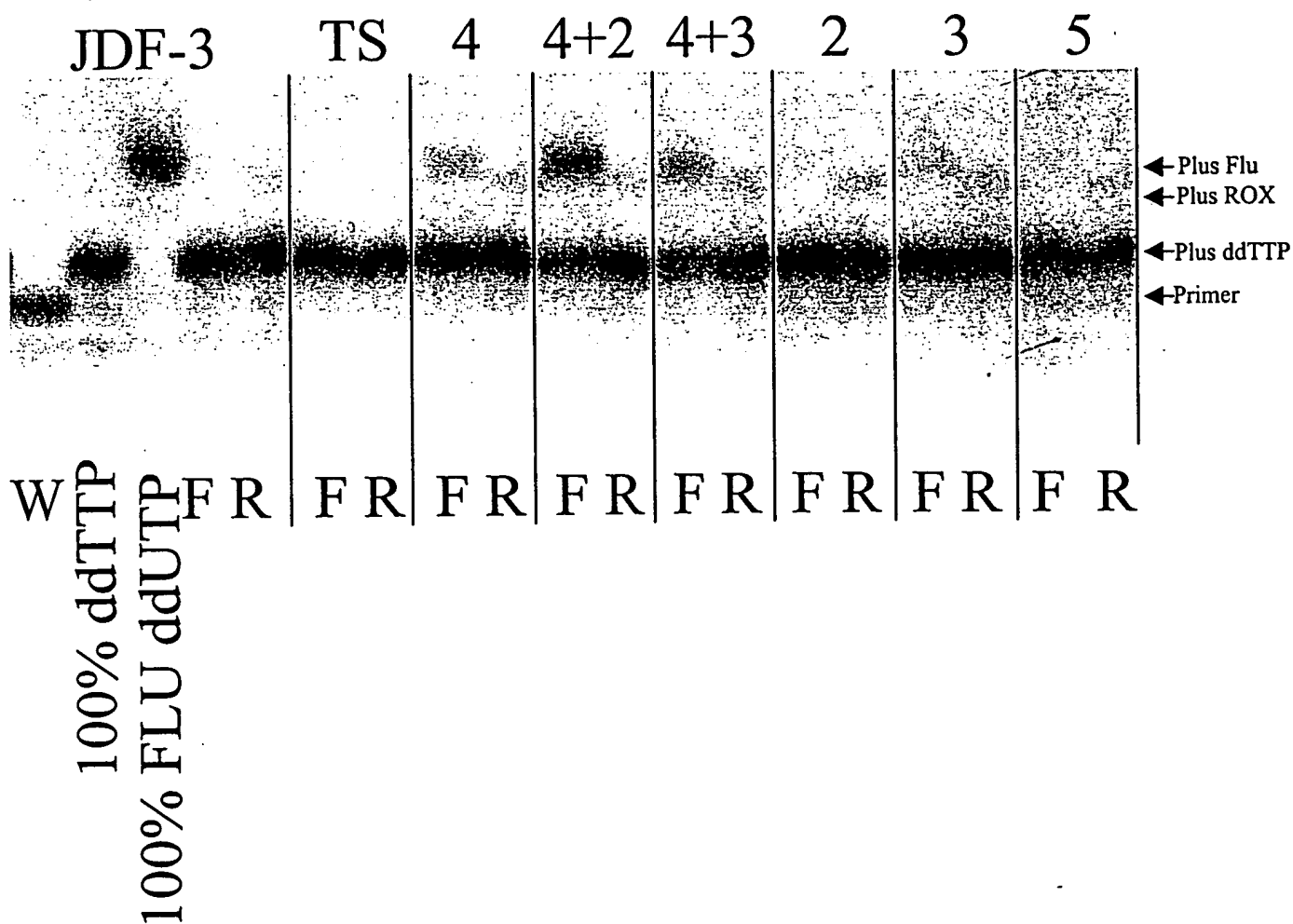


Figure 12

Flu ddUTP signal/ddTTP signal

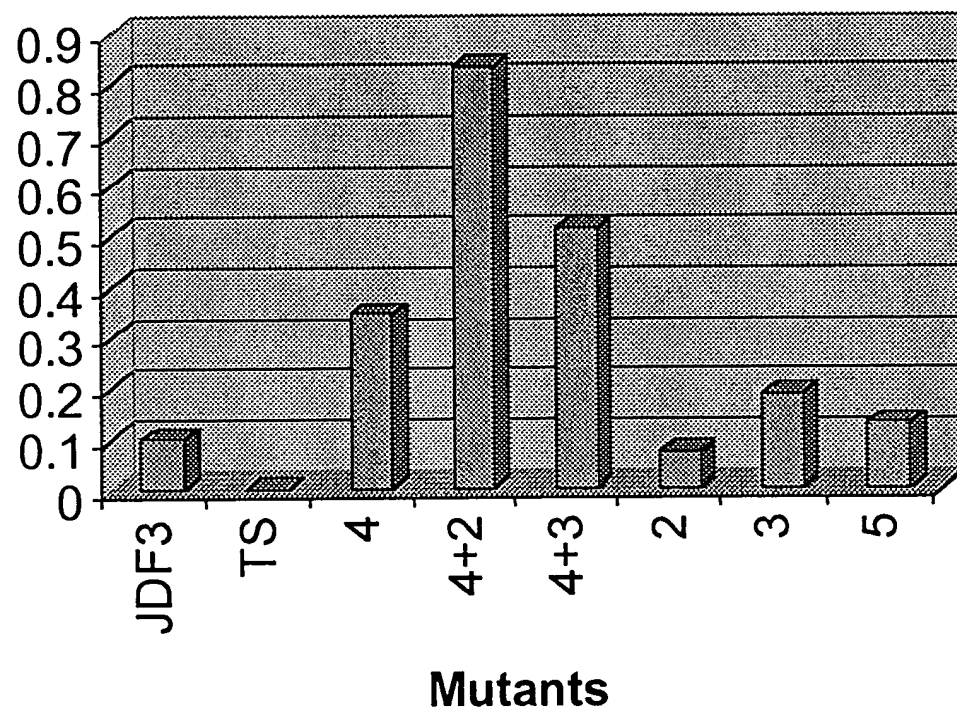


Figure 13

4	1	-----LVXNAXSTGNLVEWFLLRK
10	1	-----VWDVSRSTGNLVERFLLRK
13	1	-----VWDVSRSTGNLVEWFLLRK
16	1	-----VWDVSRSTGNLVEWFLLRK
18	1	-----VWDVSRSTGNLVEWFLLRK
19	1	-----VWDVSRSTGNLVEWFLLRK
28	1	-----VWDVSRSTGNLVEWFLLRK
34	1	-----VWDVSRSTGNLVEWFLLRK
41	1	-----VWDVSRSTGNLVEWFLLRK
33	1	-----VWDVSRSTGNLVEWFLLRK
48	1	-----YWSXPXLRTGNLVEWFLLRK
55	1	-----VIGTXPRSTGNLVEWFLLRK
64	1	-----XXXFWWDVSRSTGNLVEWFLLRK
Jdf3	301	TGEGLERVARYSMEDARVTYELGREFFPMEAQLSRLIGQGVWDVSRSTGNLVEWFLLRK

4	20	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
10	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
13	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
16	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
18	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
19	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
28	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
34	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
41	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
33	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
48	21	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
55	22	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
64	24	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP
Jdf3	361	AYERNELAPNKPDERELARRRGYAGGYVKEPERGLWDNIVYLDFRSLYPSIIITHNVSP

4	80	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
10	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
13	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
16	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
18	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
19	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
28	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
34	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
41	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
33	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
48	81	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
55	82	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
64	84	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD
Jdf3	421	DTLNREGCRSYDVAPEVGHKFCKDFPGFIPSLGNLLEERQKIKRKMKATLDPLEKNLLD

Figure 14

4	140	YRQRAIKILANSYYGYGCGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
10	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
13	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
16	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
18	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
19	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
28	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
34	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
41	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
33	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
48	141	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
55	142	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
64	144	YRQRAIKILANSYYGNYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD
Jdf3	481	YRQRAIKILANSYYGYGYGYARARWYCRECAESVTAWGREYIEMVIRELEEKFGFKVLYAD

4	200	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
10	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
13	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
16	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
18	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
19	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
28	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
34	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
41	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
33	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
48	201	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
55	202	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
64	204	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE
Jdf3	541	TDGLHATIPGADAETVKKKAMEFLNYINPKLPGLLELEYEGFYVRGFFVTKKKYAVIDEE

4	260	GKITTRGLEIVRRDWSEIAKETQARVLEAVLRHGDVEEAVRIVREVTEKLSKYEVPPEKL
10	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
13	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
16	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
18	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
19	261	GKITTRGLEIVRRDWSIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
28	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
34	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
41	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
33	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
48	261	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
55	262	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
64	264	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL
Jdf3	601	GKITTRGLEIVRRDWSEIAKETQARVLEAILRHGDVEEAVRIVREVTEKLSKYEVPPEKL

Figure 15